

```
!pip install bing-image-downloader
```

```
Collecting bing-image-downloader
  Downloading bing_image_downloader-1.1.2-py3-none-any.whl (5.9 kB)
Installing collected packages: bing-image-downloader
Successfully installed bing-image-downloader-1.1.2
```

```
from bing_image_downloader import downloader
```

```
# Specify a search query that includes keywords related to "FMCG waste" and text
query = "FMCG waste with text"
limit = 100 # You can adjust this number as needed
```

```
# Download images using the downloader
downloader.download(query, limit=limit, output_dir='your_output_directory')
```

```
📁 [%] Downloading Images to /content/your_output_directory/FMCG waste with text
```

```
[!!]Indexing page: 1
```

```
[%] Indexed 35 Images on Page 1.
```

```
=====
```

```
[%] Downloading Image #1 from https://www.packaging-360.com/wp-content/uploads/2019/07/folienabfall-1.jpg
[%] File Downloaded !
```

```
[%] Downloading Image #2 from https://www.startus-insights.com/wp-content/uploads/2021/09/FMCG-Waste-Management-Startups-FMCG-Sh
[%] File Downloaded !
```

```
[%] Downloading Image #3 from https://www.startus-insights.com/wp-content/uploads/2021/09/FMCG-Waste-Management-Startups-FMCG-He
[%] File Downloaded !
```

```
[%] Downloading Image #4 from http://inside-packaging.nridigital.com/inside-packaging/packaging\_sep19/what\_does\_waste-free\_mean\_
[%] File Downloaded !
```

```
[%] Downloading Image #5 from https://www.fluiditconsulting.co.uk/wp-content/uploads/2020/03/fmcg.png
[%] File Downloaded !
```

```
[%] Downloading Image #6 from https://thedrum-media.imgix.net/thedrum-prod/s3/news/tmp/90538/untitled\_design\_-\_2019-08-07t110013
[%] File Downloaded !
```

```
[%] Downloading Image #7 from https://changestarted.com/wp-content/uploads/2019/11/FMCG-Companies-on-Plastic-Waste-1-740x365.png
[%] File Downloaded !
```

```
[%] Downloading Image #8 from https://www.musclegenesis.com.au/wp-content/uploads/sites/88/2020/10/Recycling-bins-768x564.jpg
[%] File Downloaded !
```

```
[%] Downloading Image #9 from https://www.musclegenesis.com.au/wp-content/uploads/sites/88/2020/10/plastic-bottles-115069\_1280-3
[%] File Downloaded !
```

```
[%] Downloading Image #10 from https://starprintvn.com/wp-content/uploads/2022/01/FMCG\_packaging\_3-768x433.png
[%] File Downloaded !
```

```
[%] Downloading Image #11 from https://www.recycleright.ae/wp-content/uploads/2019/08/special-waste-1200x800.jpg
[%] File Downloaded !
```

```
[%] Downloading Image #12 from https://starprintvn.com/wp-content/uploads/2022/01/FMCG\_packaging\_4-min-768x433.png
[%] File Downloaded !
```

```
[%] Downloading Image #13 from https://changestarted.com/wp-content/uploads/2019/11/Fact-1-01-01-1-768x768.png
[%] File Downloaded !
```

```
[%] Downloading Image #14 from https://akm-img-a-in.tosshub.com/sites/btmt/images/stories/collag\_660\_210820092648.jpg
[%] File Downloaded !
```

```
[%] Downloading Image #15 from https://is3-ssl.mzstatic.com/image/thumb/Purple118/v4/65/a9/fd/65a9fdf7-e96a-9c33-5dcf-c1f4c9d6ac
[%] File Downloaded !
```

```
[%] Downloading Image #16 from https://thumbs.dreamstime.com/z/zero-waste-text-cut-old-textile-background-photo-inscription-whit
[%] File Downloaded !
```

```
!pip install pytesseract
```

```
Collecting pytesseract
  Downloading pytesseract-0.3.10-py3-none-any.whl (14 kB)
Requirement already satisfied: packaging>=21.3 in /usr/local/lib/python3.10/dist-packages (from pytesseract) (23.1)
Requirement already satisfied: Pillow>=8.0.0 in /usr/local/lib/python3.10/dist-packages (from pytesseract) (9.4.0)
Installing collected packages: pytesseract
Successfully installed pytesseract-0.3.10
```

```
!pip install Pillow
```

```
Requirement already satisfied: Pillow in /usr/local/lib/python3.10/dist-packages (9.4.0)
```

```
!apt install tesseract-ocr
```

```
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  tesseract-ocr-eng tesseract-ocr-osd
The following NEW packages will be installed:
  tesseract-ocr tesseract-ocr-eng tesseract-ocr-osd
0 upgraded, 3 newly installed, 0 to remove and 16 not upgraded.
Need to get 4,816 kB of archives.
After this operation, 15.6 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tesseract-ocr-eng all 1:4.00~git30-7274cfa-1.1 [1,591 kB]
Get:2 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tesseract-ocr-osd all 1:4.00~git30-7274cfa-1.1 [2,990 kB]
Get:3 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tesseract-ocr amd64 4.1.1-2.1build1 [236 kB]
Fetched 4,816 kB in 1s (3,541 kB/s)
Selecting previously unselected package tesseract-ocr-eng.
(Reading database ... 120901 files and directories currently installed.)
Preparing to unpack .../tesseract-ocr-eng_1%3a4.00~git30-7274cfa-1.1_all.deb ...
Unpacking tesseract-ocr-eng (1:4.00~git30-7274cfa-1.1) ...
Selecting previously unselected package tesseract-ocr-osd.
Preparing to unpack .../tesseract-ocr-osd_1%3a4.00~git30-7274cfa-1.1_all.deb ...
Unpacking tesseract-ocr-osd (1:4.00~git30-7274cfa-1.1) ...
Selecting previously unselected package tesseract-ocr.
Preparing to unpack .../tesseract-ocr_4.1.1-2.1build1_amd64.deb ...
Unpacking tesseract-ocr (4.1.1-2.1build1) ...
Setting up tesseract-ocr-eng (1:4.00~git30-7274cfa-1.1) ...
Setting up tesseract-ocr-osd (1:4.00~git30-7274cfa-1.1) ...
Setting up tesseract-ocr (4.1.1-2.1build1) ...
Processing triggers for man-db (2.10.2-1) ...
```

```
import pytesseract
from PIL import Image
```

```
# Get the folder path
folder_path = r"C:\Pictures"
```

```
# Create a text file
text_file = open("output.txt", "w")
text_file

<_io.TextIOWrapper name='output.txt' mode='w' encoding='UTF-8'>
```

```
import os
from PIL import Image
import pytesseract
```

```
folder_path = r"/content/your_output_directory/FMCG waste with text"
text_file = open("output.txt", "w")
```

```
# Iterate over the images in the folder
for i in range(1, 101):
    for extension in ["jpg", "png"]:
        image_path = f"{folder_path}/Image_{i}.{extension}"
```

```
    # Check if the image file exists
    if os.path.exists(image_path):
        # Load the image
        image = Image.open(image_path)

        # Convert the image to text
        text = pytesseract.image_to_string(image)

        # Write the text to the text file
        text_file.write(text + "\n")
```

```
# close the file
text_file.close()
```

